How early L2 children perform on Italian clinical markers of SLI

ACCEPTED IN: CLINICAL LINGUISTICS AND PHONETICS

TITLE

How early L2 children perform on Italian clinical markers of SLI: a study of clitic production and nonword repetition

Maria Vender[a], Maria Garraffa[b], Antonella Sorace[c], Maria Teresa Guasti[d]

[a] University of Verona

[b] Heriot-Watt University

[c] University of Edinburgh

[d] University of Milano-Bicocca

Abstract

Early Second Language (EL2) learners generally perform more poorly than monolinguals in specific language domains, presenting similarities with children affected by Specific Language Impairment (SLI). As a consequence, it can be difficult to correctly diagnose this disorder in EL2 children. The current study investigated the performance of 120 EL2 and 40 age-matched monolingual children in object clitic production and non-word repetition, which are two sensitive clinical markers of SLI in Italian. Results show that EL2 children underperform in comparison to monolinguals in the clitic task. However, in contrast to what is reported on Italian-speaking children with SLI, EL2 children tend not to omit clitics but instead produce the incorrect form, committing agreement errors. No differences are found between EL2 and monolingual children on non-word repetition. These results suggest that, at least in Italian, EL2 children only superficially resemble children with SLI and, on closer inspection, present a qualitatively and quantitatively different linguistic profile.

Keywords: bilingualism, Specific Language Impairment, clitic production, nonword repetition

Introduction

One of the most interesting challenges posed by bilingualism concerns the identification of language impairments in bilingual children and early second language learners (EL2). In fact, bilingual and EL2 children often perform more poorly in comparison to their monolingual peers in specific language domains. A number of studies have shown that bilinguals have a smaller vocabulary in both languages than monolinguals and underperform in standardized receptive vocabulary tests (Oller et al. 2007, Bialystok et al. 2010). Moreover, weaknesses have been found in the domain of morphosyntax, especially in those tasks, which impose high processing costs (Serratrice et al. 2004, Sorace et al. 2009). The attested presence of lexical and morphosyntactic difficulties in bilingual and EL2 children may invite to draw a parallel with children suffering from Specific Language Impairment (SLI). Furthermore, the increased attention to the recognition of disabilities and of the importance of providing prompt services to affected children are leading communities to seek information about how to identify SLI in bilingual environments.

SLI is a neurodevelopmental disorder affecting approximately 7% of preschool children and characterized by language abilities below age expectations, despite normal cognitive abilities and absence of physical and neurological deficits (Leonard, 1998; Rice 2004). Children with SLI present a delay in their language which does not get completely resolved over time and display deficits in comparison to their typically developing peers both in the lexical and in the morphosyntactic domain, domains that are temporarily compromised also in EL2 children. Bilinguals and children with SLI have been found to perform superficially similarly in morphpsyntactic tasks across different languages. Both populations display lower accuracy rates in comparison to unaffected monolinguals in the production of tense and nontense morphemes, (Paradis & Crago 2000), in the ability to detect ungrammatical uses of these

morphemes (Paradis et al. 2008), in the production of direct object clitics (Grüter 2005) and in the acquisition of word order structures (Håkanson & Nettelbladt 1993, Håkanson 2001).

The presence of these similarities can have an impact on the diagnosis of SLI in bilingual children, resulting both in the over-diagnosing and the under-diagnosing of the impairment, due to the absence of diagnostic tools expressly designed for the identification of language impairments in bilinguals and to the limited normative data concerning the trajectory of EL2 acquisition (Bedore & Pena 2008).

One way to tackle the problem is to examine the proficiency of bilingual/EL2 children with the clinical markers of SLI in a given language; that is in those areas which are particularly vulnerable for children with SLI. For what concerns Italian, the production of 3rd person direct object clitic pronouns and the repetition of nonwords are considered two clinical markers which permit to accurately distinguish children with SLI from age-matched typically developing children. In this framework, it would be crucial to analyse how EL2 children perform with both markers, to determine whether this line of investigation is worth pursuing for the purpose of identifying SLI in EL2 children.

Clitic pronouns acquisition: monolingual children, children with SLI and L2 learners of Italian

The production of direct-object (DO) clitic pronouns in Italian involves the integration of phonological, morphological, syntactic and pragmatic information, which requires sophisticated linguistic competence and efficient processing abilities. Monolingual Italian children with a typical linguistic development normally start to optionally produce DO clitics around the age of two years. When they produce DO clitics, they do not display placement errors and do not mix with tonic pronouns (Guasti 1993/1994; Schaeffer 2000; Caprin & Guasti 2009; Moscati & Tedeschi 2009). Agreement errors are also occasionally observed at age 3-4

(Tedeschi 2006); as reported by Pizzuto and Caselli (1992) the feminine singular *la* is acquired first, followed by the masculine singular *lo*. When children produce an incorrect clitic, they generally show an overall preference for *lo*, suggesting it is the unmarked clitic form in Italian, a sort of protoclitic which may not actually carry the grammatical features of the clitic but rather act as a grammatical placeholder (Leonard & Dispaldro 2013; Dispaldro et al. 2009). Nonetheless, young typically developing native speakers of Italian show an optional use of clitic pronouns, variably omitting them usually up to age 4, with gradual improvements as they grow up.

Interestingly, the period of optional use of clitics is significantly extended in children with SLI, who manifest a persistent tendency to omit them, producing sentences which lack the internal argument and are therefore ungrammatical in Italian. Bortolini and colleagues (Bortolini et al. 2002, 2006) tested normally developing and children with SLI aged between 3;7 and 5;5 years in an elicitation task in which participants were shown two drawings and prompted to complete a sentence like "*Qui la bambina compra il gelato e qui…"* ('Here the girl buys the ice-cream, and here…") where the target answer should be "*lo mangia*" ("she eats it"). In the Bortolini et al. studies, the authors compared the performance of the groups with respect to the production of target clitics, omissions of the pronoun, and clitic substitution, when an incorrect DO clitic was produced. There were no instances in which the full NP was uttered instead of the clitic. The authors found that unimpaired children produced the target sentence in 91% of cases, whereas children with SLI produced the correct clitics only in 18% of cases. Errors with clitics always took the form of omission in both groups of children, except for a single case of clitic substitution made by a child with SLI. Children with SLI underperformed also in comparison to younger children aged between 2;10 and 4;0 and matched for Mean Length of Utterance, who produced target sentences in 72% of cases.

Difficulties in the production of clitics have been also reported in another study conducted by Leonard and Dispaldro (2013), who found that Italian-speaking preschool children with SLI (mean age 4;9) are even more impaired in clitic production than 18 months younger typically developing children. In their study, the authors manipulated the demands of the task by introducing a syntactic priming condition in which subjects could benefit from a preceding sentence prime. In their task children were asked to answer a sentence eliciting a clitic, such as *Cosa succede alla televisione?* ('What is happening to the television?'), where the target utterance was *Mowgli la spegne* ('Mowgli turns it off'). Children with SLI and control children were compared in their production of target clitics, omissions, and clitic substitutions. The production of full NPs and irrelevant sentences (e.g. "I don't know") were treated as unscoreable and not reported in the paper. Results showed that children with SLI tend to omit the clitic in the control condition much more often than unimpaired subjects, as in the (2006) study by Bortolini and colleagues, whereas clitic substitution is more common in the priming condition, where children with SLI show an overuse of the clitic *lo* as a substitute. The authors argue that the facilitation effect generated by the priming condition prompts impaired children to produce a clitic form, though often incorrect, whereas omissions increase when the sentence demands are greatest, as in the control condition.

Arosio et al. (2014) confirmed that clitic production is still challenging for 7 year-old children with SLI, even though the typology of errors committed is different. Participants in their study were shown some pictures on a computer screen and told a short story about one character performing an action; they were then asked to answer a question about this story eliciting a clitic pronoun, such as *Cosa fa il bambino alla farfalla?* ('What does the boy do to the butterfly?), where the target utterance was *La prende* ('he catches it'). As in Bortolini et al. (2006) and Leonard and Dispaldro (2013), the authors compared group performances with respect to the production of target clitics, clitic substitutions, omissions and production of NPs.

Moreover, they also considered the production of indirect clitics, where an indirect clitic was used instead of a DO clitic, and the production of irrelevant sentences. Results revealed that even school-aged children with SLI produce a lower number of target clitics in comparison to unimpaired subjects; interestingly, their most common error is the production of a nominal phrase (NP) instead of the clitic, which is not felicitous in the target context. No significant differences were found between the two groups in the rate of omissions, production of wrong clitics, indirect clitics and irrelevant sentences.

The production of DO clitic pronouns is also challenging for adult L2 learners of Italian (Leonini 2006). When adult L2 learners do not produce clitics, they usually replace them with their corresponding NP; clitic case morphology also appears to be problematic for this population (Belletti & Hamann 2004, Santoro 2007).

In contrast, the acquisition of reflexive (RE) clitics is not problematic for children with SLI or for typically developing monolingual children. RE clitics are commonly produced by monolingual children from age 2 in an adult-like fashion and they are used much more consistently than DO clitics (Snyder et. al, 1995; Caprin & Guasti, 2009). This asymmetry between DO and RE has also been reported in French (Zesiger et al. 2010). The production of RE clitics is unproblematic for school age children with SLI, as reported by Arosio and colleagues (2014).

To conclude, the production of clitic pronouns is challenging for both children with SLI and L2 learners of Italian, whereas RE clitics are acquired and mastered earlier by both impaired and unimpaired children.

Nonword repetition

In nonword repetition tasks subjects are asked to repeat meaningless but pronounceable words, modelled according to the phonotactic rules of their native language. This kind of test

taps the ability to store and rehearse verbal information in short term memory and provides a particularly 'clean' measure of phonological memory and phonological awareness.

Nonword repetition is sensitive to different language disorders: low accuracy rates in nonword repetition have been found in dyslexic children (Vender to appear, Snowling 1981, Roodenrys & Stokes 2001) and in children with SLI. Performance on nonword repetition is considered a clinical marker of SLI across different languages, including Italian (Casalini et al. 2007; Vernice et al. 2013, Bishop et al. 1996; Conti-Ramsden 2003; see also Coady & Evans 2008 for an extensive review). Bortolini and colleagues (2006) showed that nonword repetition is a reliable clinical marker for SLI in Italian, reporting that preschool children with SLI are significantly more impaired than their peers in this task. Specifically, their accuracy decreased as a function of the increasing length, in terms of number of syllables, of the word to repeat. As for the scoring method, the authors considered the nonword correct if all consonant and vowel segments were produced correctly and there were no additions, whereas they scored as errors all substitutions, omissions and additions. Analysing the data, the authors found that SLI children consistently omitted entire syllables, more precisely non-final weak syllables in long nonwords, whereas control children showed a significantly greater accuracy and did not generally omit syllables. Specifically, children with SLI repeated correctly a significantly smaller percentage of non-words (mean 40.25, SD 37.59) compared to control children (mean 79.89, SD 19.85). The discrepancy observed between SLI and typically developing children confirmed that nonword repetition is a reliable clinical marker for Italian. Similarly, Casalini and colleagues (2007) report that Italian SLI children are impaired in the repetition of both words and nonwords in comparison to age-matched controls. Subjects were asked to repeat three sets of 20 stimuli each: words, nonwords built with non-existing morphemes and nonwords composed by a root and an existing derivational suffix. The author compared the performances of children with SLI and typically developing children, reporting that language-

impaired children were less accurate in all three measures in comparison to controls, and therefore confirming that the impairment in repetition tasks can be considered as a marker of SLI.

A number of studies reported that typically developing bilingual children perform significantly better than monolingual children with SLI but less optimally than monolingual unimpaired controls in nonword repetition tasks administered in English (Girbau and Schwartz 2008; Khonert, Windsor and Kim 2006; Windsor et al. 2010) and Dutch (Messer and colleagues 2010). Conversely, Guasti et al. (2013) did not find differences between preschool EL2 children and age-matched monolinguals in nonword repetition in Italian. The partial discrepancy between these results can be explained on the one hand by observing the phonological structures of the languages tested and, on the other hand, the age of exposure to the L2. English and Dutch have a more complex phonology in comparison to Italian, which has a lower number of syllable types and of consonantal clusters. Accordingly, nonword repetition appears to be less demanding for those children whose L2 has a less complex phonotactic structure, like Italian. A similar conclusion was drawn by a study by Tamburelli et al. (2014), which showed that Polish-English bilingual children have an advantage over English monolinguals in the acquisition of English phonological structure. The authors ascribe this benefit to the higher phonological complexity of Polish in comparison to English, proposing that the exposure to complex phonological structures can accelerate the development of less phonologically complex languages.

Given this background, the research question of this study is: how do typically developing L2 children of Italian perform with respect to DO clitics and non-word repetition? The answer to this question is crucial in order to gain information about typical L2 in areas that are weak in monolingual children with SLI.

Methods

Participants

The experimental task was administered to 120 early L2 (EL2) preschool children who were acquiring Italian as their L2. They were divided in three distinct groups according to their L1: 40 Albanian-speaking L1 children (ALB; mean age 59 months, SD= 8.43), 40 Arabic-speaking L1 children (ARA; mean age 57 months, SD=6,14) and 40 Romanian-speaking L1 children (RUME; mean age 58 months, SD = 7,42). All EL2 children were exposed to their L1 from birth and had at least one year of exposure to Italian. Age of first exposure to Italian varied significantly: the majority of children were exposed to Italian when they entered daycare centers (around 30-42 months of age), but a number of children were exposed also to Italian from birth, especially those who had older brothers and sisters speaking Italian at home. The control group was composed by 40 monolingual Italian children matched for chronological age and sex (MON; mean age 58 months, SD = 6.36). A one-way ANOVA was carried out on these data revealing that there were no significant differences in the age of the subjects ($F(3, 152) = .794$, $p= .499$). All children were recruited from public kindergartens in the area of Trento (Italy).

Materials

All children were administered a set of preliminary tasks in order to ensure cognitive comparability and to have a more precise picture of their linguistic competence in Italian. Subsequently children were tested in clitic elicitation and in nonword repetition tests. A description of the materials used follows below.

Preliminary Tasks

Raven test. In order to guarantee comparability in nonverbal cognitive ability, all participants were tested in the standardized Raven's Coloured Progressive Matrices test (Raven et al. 1998).

Bilingual Language Exposure Questionnaire. Detailed information about EL2 children's exposures to Italian were collected administering a version of the questionnaire Utrecht Bilingual Language Exposure Calculator (UBiLEC) (Unsworth et al. 2012) that we adapted to Italian. The questionnaire, completed by parents, provides a detailed description about children's exposure to Italian. We collected information about the children's age of irst exposure (AFE) to Italian, their current quantity of exposure (QE) to the L2, the raditional length of exposure (TLE), which is calculated as the child's chronological age minus their age at first exposure to Italian, and the cumulative length of exposure (CLE), which is a composite measure that considers other variables to determine the actual exposure to the L2 over time.

PPVT-R. More precise information about the participants' linguistic competence in Italian was collected administering a receptive standardized vocabulary test (PPVT-R, Peabody Picture Vocabulary Test – Revised, Stella et al. 2000).

Comprendo. The children's comprehension of Italian was assessed using a subset of the test Comprendo (Cecchetto et al. 2012), a picture-selection task in which the child has to point to a picture from an array of four that match a sentence uttered by the experimenter. Our selection of sentences examined children's understanding of 7 types of linguistic structures (active sentences, passive sentences, dative sentences, subject and object peripheral relative clauses, coordinated sentences and verbal inflections) with 3 items per structure for a total of 21 items.

Clitic production task. Production of DO clitic pronouns was examined through an elicitation task, developed on the basis of the task administered by Arosio et al. (2014) and similar to that used in Leonard and Dispaldro (2013). During the task, subjects were shown some pictures displayed on a computer screen and told a short story that always involved two or three characters performing an action. Descriptions were digitally recorded by a feminine Italian native speaker and played through loudspeakers connected to the laptop. When the first picture appeared, the characters of the story were introduced and the child was told that one character wanted to perform some action to the other/s. In the second picture, portraying the character performing that action, the child was asked to answer a question eliciting the DO clitic pronoun about what the character did. An example of the task is reported below.

(1) Experimenter: "In questa storia ci sono un nonno e una bambina. La bambina vuole baciare il nonno. Guarda adesso cosa succede. Cosa fa la bambina al nonno?"
('In this story there are a grandfather and a girl. The girl wants to kiss the grandfather. Look at what is happening now. What does the girl do to the grandfather?')
Target answer: *Lo bacia* ("she kisses him")

We elicited 12 sentences containing a third person DO clitic: (i) 3 with the masculine singular pronoun "*lo*", (ii) 3 with the feminine singular "*la*", (iii) 3 with the masculine plural "*li*" and (iv) 3 with the feminine plural "*le*". The task was preceded by a familiarization section consisting of five training items in which participants were invited, if necessary, to answer the questions producing the clitic pronoun. The experimental items were intertwined with four items eliciting the production of the third person singular reflexive (RE) clitic pronoun *si* ("itself"). An example is reported below:

(2) In questa storia ci sono una bambina e una signora. La bambina vuole asciugarsi. Guarda cosa succede adesso. Cosa fa la bambina?

("In this story there are a girl and a lady. The girl wants to dry herself. Look at what is happening now. What does the girl do?")

Target answer: *Si asciuga* ("she dries herself")

All verbs used were conjugated in the present tense and were obligatory transitive.

Nonword repetition task. EL2 children's nonword repetition skills were tested by administering a standardized nonword repetition test (NWR, Cornoldi et al. 2009). In this task, the subject was asked to listen carefully to a nonword pronounced by the experimenter and then to repeat it. The NWR test included 25 stimulus of increasing length and complexity, ranging from one to four syllables. There were 5 nonword items for each length, with length 2 occurring twice. All stimuli conformed to Italian phonotactic patterns; the accent of three and four syllables nonwords was either on the penultimate syllable (unmarked stressed pattern in Italian), or on the antepenultimate or on the initial syllable (marked stressed pattern).

Each test session was preceded by a familiarization session with two training items. The subject's score was calculated considering the total number of syllables correctly repeated, for a maximum of 60 syllables. As in Bortolini et al. (2006), omissions and addition of sound and syllables were considered errors, as well as repetition of syllables with incorrect sounds and simplifications of consonant clusters.

Procedure

Each child was individually tested in a quiet room; the experimental protocol was administered in two separate sessions lasting about 30-40 minutes each. Each session was registered and all materials were transcribed and re-examined by independent researchers.

Results

Statistical Analysis. Statistical analysis of measures of exposure were performed using ANOVA with Group as between-subject factor (henceforth BS), whereas statistical analyses of Raven, PPVT-R, Comprendo, Nonword repetition score were performed using ANOVA with Group (MONO, ALB, ARA, RUME) as between-subject factor and cumulative length of exposure, traditional length of exposure and quantity of exposure as covariates. Significant main effects were followed up using Bonferroni's post-hoc comparisons. Significant values are always meant to be less than 0.05. We report only significant main effects, indicating partial eta squared (η^2_p) as a measure of effect size. Given that measures of clitic production (response accuracy) are categorical, we used mixed logit models (R Development Core Team, 2011), employing a stepwise forward inclusion procedure and starting with a model without factors. Then, we added one predictor at a time and compared a model including the predictor against one without it, using a χ-square test (Jaeger, 2008). Following this procedure, we established which factors contributed to the model's fit. Next, based on *z* values (Wald statistics), we obtained an estimate of the statistical significance of each predictor in the model. Both first-level effects and interactions between the fixed-effect factors were examined. Group, Comprendo, PPVT-R, nonword and Raven scores, cumulative length of exposure (CLE), traditional length of exposure (TLE) and quantity of exposure (QE) were introduced as potentially significant fixed effect. In all models, monolingual children were the reference category for the predictor Group. Subjects and items were introduced as random factors.

Preliminary Tasks.

Out of the 160 subjects who took part in the experiment, four of the children (1 ARA, 2 RUME, 1 MON) were excluded since they scored 1.5 SD below the mean score for their age in the Raven test. Moreover, two of the ARA children were not able to perform the clitic production task and were therefore excluded from the sample. Descriptive statistics concerning Group data in the preliminary tasks and the nonword task are reported in Table 1.

INSERT TABLE 1 HERE

As it can be noted from the table, all the EL2 were first exposed to Italian at the same age (AFE). However, ARA children had a lower quantity of exposure (QE) in comparison to the other two groups of EL2 children, resulting in a lower cumulative length of exposure to the L2 (CLE). A one way ANOVA evidenced that there is a significant difference among the three groups of EL2 regarding QE ($F (2, 112)= 4.5$ $p= .01$, $\eta^2_p=0.07$), due to ARA having been less exposed to Italian than ALB ($p=.01$). CLE was marginally significant ($F (2, 112) = 2.81$ $p= .06$, $\eta^2_p=0.04$), again due to lower exposure to Italian of ARA in comparison to ALB children ($p=.06$). No other differences were found concerning the AFE, CLE and TLE. These data suggest that, even though all three groups of EL2 children were exposed to Italian at the same age, ARA received less exposure to Italian.

As for the preliminary tasks, results from a one-way ANOVAs revealed no main effect for Raven scores, whereas a main effect of Group for PPVT-R raw scores was found ($F (3, 150)= 16.4$, $p<.01$, $\eta^2_p=0.24$), showing that EL2 children are weaker than monolinguals in receptive vocabulary. Multiple post-hoc comparison showed that MON achieved higher scores than all groups of EL2 ($p<.01$), and that ARA were weaker than RUME ($p<.01$). As for Comprendo, Group ($F(3,150)= 5,41$, $p<.01$, $\eta^2_p=0.10$) turned out to be significant and post-hoc analyses showed that ARA differ from MON ($p<.01$) and RUME ($p<.05$). Summarizing, all

three groups of EL2 children obtained lower performance than monolingual control in the PPVT-R, with ARA performing particularly poorly, whereas only ARA underperformed in the comprehension tasks, presumably due to their shorter cumulative exposure to Italian over time.

Clitic Production Task. Children's utterances were analyzed with a rigorous scoring method and classified in 5 different categories, following Arosio et al. (2014). We considered *Target* those utterances in which the correct DO clitic was produced, regardless of the verb and tense used (e.g. *lo bacia* ["she kisses him"]. We also included in this category grammatical sentences containing an indirect object pronoun instead of a DO clitic (e.g. *Gli dà un bacio* ["she gives him a kiss"]). We classified as *Incorrect Clitic* those utterances in which an incorrect DO clitic was produced, including both gender and number agreement errors (e.g. producing *lo* instead of *la* or *li*). *NP* utterances were those answers in which the subject used a nominal phrase instead of a clitic, producing a pragmatically infelicitous answer (e.g. *bacia il nonno* ["kisses the grandfather"] instead of *lo bacia* ["kisses him"]). The *Omission* category included the answers in which the clitic was omitted, leading to an ungrammatical sentence (e.g. *bacia* ["she kisses"]). Finally, we treated as *Other* responses all the irrelevant sentences produced by the participants (e.g. *Il nonno è felice* ["the grandfather is happy']). The results of the test are reported in Table 2.

INSERT TABLE 2 HERE

As it can be immediately noted, MON produced more target structures (73%) than all three groups of EL2 children, with RUME (51%) and ALB (48%) performing better than ARA (29%). A very common error for both monolingual and EL2 children was the use of an incorrect clitic (33% ALB, 29% RUME, 24% ARA, 16% MON). Considering the total number of clitic

produced, both correct and incorrect, the two groups of EL2 produced more than 80% of the clitics (81% for ALB and 80% for RUME), approaching monolingual children with 89% (16% of incorrect forms). The behavior of ARA children was different, as their production of incorrect clitics (24%) was lower as well as their total production of clitics (53%); their most common error was the production of irrelevant sentences (28%), which significantly distinguished their performance from that by the other groups of children (9% ALB, 6% RUME, 2% MON). Interestingly, omissions were very low in all groups (4% ALB, 9% ARA, 8% RUME, 3% MON), as well as productions of NPs (6% ALB, 9% ARA, 6% RUME, 6% MON).

These observations are confirmed by the statistical analysis reported in Table 3. Selecting *Target* response as the dependent variable, *Group* [$\chi^2(3) = 48.74$ $p<.001$] and *Comprendo* [$\chi^2(1) = 24.73$ $p<.0001$] contributed significant information, whereas *PPVT-R*, *Nonword* and *Raven* scores did not. Table 4 indicates that the probability of producing a target clitic decreases in all three groups of EL2 children (negative coefficients from the reference category, MON, to the contrasting category, EL2 children). By using ARA as reference category, we found that this group produced fewer target clitics than the other two EL2 groups (positive coefficients from ARA to the other groups).

Using *Incorrect Clitic* as dependent variable we found that only Group [$\chi2(3) = 14.45$ $p<.01$] added significant information to the model, indicating that ALB and RUME produced more wrong clitics than MON (positive coefficient from the reference category, MON, to the contrasting category, Groups of EL2 children). By changing the reference categories, we compared each group of children with the others and found that ARA produced a lower number of incorrect clitics in comparison to ALB.

Selecting *Other* responses as dependent variable we found that *Group* [$\chi2(3) = 37.24$ $p<.001$], *Comprendo* [$\chi2(1) = 11.6$, $p<.001$] and *PPVT-R* [$\chi2(1) = 4.05$, $p<.05$] contributed to

the model's fit. ARA produced more irrelevant utterances than MON and the other two groups of EL2 children. No difference was found between MON, on the one hand, and RUME and ALB, on the other. In the models analyzing *Omissions* and *NP* responses neither *Group* nor the other independent variables contributed to the model's fit.

INSERT TABLE 3 HERE

Subsequently, we examined whether the EL2 Groups' performance was predicted by CLE, TLE and AFE, using ALB as the reference category for *Group* (see Table 4). We ran a model with *Target* as dependent variable and found that *Group* contributed significant information to the model's fit [$\chi^2(2) = 14.89$ p<.001] as did *CLE* [$\chi^2(1) = 22.35$ p<.001]. When *Other* responses was used as dependent variable *Group* provided significant information [$\chi^2(2) = 20.78$ p<.001] as did CLE [$\chi^2(1) = 7.18$ p<.001], confirming that ARA performance, characterized by a lower number of target sentences and a higher rate of irrelevant responses, is predicted by *CLE*.

INSERT TABLE 4 HERE

In order to establish whether the types of errors were different across groups, we performed an analysis of the errors committed by children. Errors were classified in four different categories: gender (e.g. *lo* for *la*, or *li* for *le* and viceversa) number (*la* for *le* or *lo* for *li* and viceversa), gender and number (e.g., *lo* or *la* for *le* or *li*, respectively) and case (*gli*, an indirect object clitic, for one of the direct object clitics). Data are reported in Table 5.

INSERT TABLE 5 HERE

ALB and RUME committed more errors in comparison to both MON and ARA; the most common error for all groups was Gender error, followed by Number error, which is also quite common for ALB and RUME. The statistical analysis revealed that *Group* added significant information to the model with *Gender* [$\chi^2(3) = 20.94$ p<.01], and *Number* [$\chi^2(3) = 614.94$ p<.01] as dependent variables, indicating that ALB and RUME substituted the target clitic with

one featuring the wrong gender or number more often than MON. No other group effect was found (see Table 6).

INSERT TABLE 6 HERE

Following Leondard & Dispaldro et al (2013) we also examined whether there was a preference for a particular clitic when clitic substitution occurred. Table 7 reports the raw number of errors divided by clitics, with each column indicating how many times a given clitic was incorrectly chosen.

INSERT TABLE 7 HERE

Clitic substitution was quite common for ALB, RUME and also MONO children, resulting in the overuse of the masculine singular *lo*, which was the most frequently chosen substitute, followed by the feminine singular *la*. ARA, in contrast, produced a lower number of substitution error and did not show a clear preference for *lo* over *la.* The overuse of *lo* resembles the tendency found in monolingual children, reviewed above. With respect to the plural clitic, all children show a clear preference for the masculine *li* as a substitute. Statistical analyses did not yield any effect of group.

Finally, we examined the use of the reflexive clitic *si*. Table 8 reports the number of *Target* clitics, of other *Incorrect Clitics*, *Omissions* and *Other* irrelevant responses. Monolingual and EL2 produced a fair number of target RE clitics, with the exception of ARA, who frequently omitted the clitic or gave irrelevant responses.

INSERT TABLE 8 HERE

This is confirmed by the statistical analysis. When *Target* was used as dependent variable *Group* [$\chi^2(3) = 60.36$ $p<.001$], *Comprendo* [$\chi^2(1) = 7.53$ $p<.01$] and *PPVT-R* [$\chi^2(1) = 4.92$ $p<.05$] contributed significant information. Table 10 indicates that the probability of producing a target RE clitic decreases in ARA with respect to MON and that performance is predicted by the score in *Comprendo* and in the *PPVT-R*. When *Omission* and *Incorrect clitic* were selected as independent variables, only *Group* contributed to the model fit with [$\chi^2(3) = 27.82$ $p<.001$] and [$\chi^2(3) = 9.95$ $p<.01$], respectively. Finally, when *Other* was the independent variable, *Group*

contributed significant information [$\chi^2(3) = 25.36$ $p<.001$], as did *Comprendo* [$\chi^2(1) = 8.54$ $p<.001$] (see Table 9).

INSERT TABLE 9 HERE

Nonword Repetition Task.

Results of the nonword repetition task are reported in Table 10. The participants' score was calculated considering the total number of syllables correctly repeated, for a maximum of 60 syllables. As in Bortolini et al. (2006), omissions and addition of sound and syllables were considered errors, as well as repetition of syllables with incorrect sounds and simplifications of consonant clusters.

INSERT TABLE 10 HERE

As is evident from the table, the four groups of children show a very similar performance. This was confirmed by the statistical analysis which failed to find significant differences among the groups.

Discussion

The present study aimed to investigate how EL2 children perform with respect to two of the clinical markers of SLI in Italian, namely the production of DO clitic pronouns and the repetition of nonwords. The performance of 40 typically developing Italian monolinguals was compared to that of 120 age-matched typically developing EL2 children, including 40 Albanian-speaking L1 children), 40 Arabic-speaking L1 children and 40 Romanian-speaking L1 children. All EL2 children came from immigrant families, had been exposed to Italian at least for 1 year and were living in the same region of the monolingual children; information about maternal education and socio-economic status were not collected, but it is reasonable to

presume that EL2 children came from a lower socio-economic background in comparison to monolinguals. However, cognitive comparability between the groups was assured by the administration of the Raven task and of the two language tasks assessing receptive vocabulary (PPVT-R) and comprehension (Comprendo).

In order to analyse the EL2 children's use of the Italian clinical markers of SLI, we administered a clitic production task and a nonword repetition task.

Results demonstrate that clitic production is challenging for EL2 preschool children who were found to perform more poorly than MON, as expected. Interestingly, the most common error committed by ALB and RUME was the production of an incorrect clitic, with an overuse of the clitic *lo*, which is also the most frequent error shown by MON. However, if we consider the total production of clitic pronouns, both correct and incorrect, the performance of ALB and RUME approached that of MON. The behavior shown by ARA was slightly different: their most common error was the production of irrelevant sentences.

The discrepancy reported between ALB and RUME, on the one side, and ARA on the other cannot simply be due to transfer from the different L1 of the children. The three languages all have a clitic pronominal system, which is inflected for person, number and gender in Italian, Arabic and Romanian and only for person and number in Albanian. The only significant differences among the four languages concerns clitic placement: clitics generally precede the verb in Italian, Albanian and to some extent in Romanian, whereas they are always enclitic in Arabic. If ARA children's greater difficulties with clitic production had really been affected by negative transfer we should have expected a higher rate of placement errors, which instead were not committed at all.

A more plausible explanation for ARA's poorer performance relates it to lower exposure to Italian and to their less developed linguistic competence in Italian, as demonstrated by lower scoring in vocabulary and comprehension tasks. As the statistical analysis suggests,

children who have a better competence in Italian are more skilled in the production of clitic pronouns, whereas children with a lower competence have more difficulties with clitics, which results in a higher production of irrelevant sentences. It is plausible to assume, then, that ARA's weaker performance was due to their lower cumulative exposure to Italian. This explanation is supported by the analysis of reflexive pronouns: differently from ALB and RUME, who produced RE clitics in a monolingual-like fashion, ARA showed marked difficulties, omitting the reflexive, producing incorrect pronouns or uttering irrelevant sentences. Moreover, their performance in RE clitic production was predicted by their linguistic competence, indicating that those children who had lower vocabulary and comprehension skills in Italian had more difficulty in producing the correct RE clitic. This result seems to support the idea that cumulative exposure of the participants to Italian, more than their L1, predicts their ability to produce correct clitics.

Summarizing, our results show that EL2 children performed more poorly in comparison to monolinguals in clitic production and that their most common error consisted in producing the incorrect clitic, in the case of ALB and RUME, or in uttering irrelevant sentences, in the case of ARA, whereas omissions and production of NPs are very infrequent and similar to those shown by monolinguals.

These results can be compared to those obtained in the studies on children with SLI reviewed above. The scoring criteria adopted our study are identical to those used in Bortolini et al. (2006) (although they did not find the whole range of responses we did) and Arosio et al. (2014). Given that Leonard and Dispaldro (2013) excluded sentences with full NPs, their results can be compared to ours only with some caution. However, the studies discussed report that at age 5 the most common error committed by SLI children is omission, whereas at age 7 it is the production of NPs (Bortolini et al. 2002, 2006; Arosio et al. 2014, Leonard and Dispaldro 2013). Comparing these results to ours, it is worth noting that omission, which is

distinctive at this age for children with SLI, is not an option for any of the groups of EL2 children we tested. The high tendency to produce incorrect clitics shown by EL2 children suggests that they are aware of the fact that clitics have to be produced in Italian, and that their difficulties are mainly limited to the choice of the correct inflection on the pronoun and related to their general competence in Italian. A reasonable prediction is that their difficulties will disappear as their competence in Italian increases. It is apparent that in preschool years clitic production is difficult for both populations, but the typology of errors committed by the two groups differentiates between them.

The second clinical marker for SLI that we examined is the repetition of nonwords. Our results demonstrate EL2 children, independently from their L1, perform as accurately as their monolingual peers, similarly to what Guasti et al. (2013) found in a smaller group of EL2. In this respect, the EL2 profile is clearly distinct from that of children with SLI, whose performance at nonword repetition is typically poor. Since nonword repetition scores provide a measure of phonological memory and phonological awareness, we can infer that EL2 do not have problems in these areas, while SLI do.

Phonological deficits are one of the explanations proposed to account for children with SLI's difficulties with clitic production. Bortolini and colleagues (2006) observed that SLI tend to omit non-final weak syllables in nonword repetition tasks and suggested that the phonological status of clitics, which are indeed non-final weak syllables, is responsible for their high rates of omissions in clitic production tasks. Their proposal has been extended by Arosio et al. (2014) who argue that phonological deficits can be one of the causes of children with SLI's poor performance, together with the morphosyntactic complexity of clitic pronouns which could exceed their processing resources.

Evidently, phonological problems cannot be held responsible for EL2 children's difficulties with clitic production, since they rarely omit clitics and their phonological memory

is unimpaired. It is more plausible to propose that their deficits, and especially their agreement errors, are due to failure in selecting the correct form due to an incompletely automatized access to lexical forms.

Therefore, it seems plausible to assume that the difficulties shown in clitic production by both SLI and EL2 children are clearly different in nature, as demonstrated by the distinct error typology in the two groups, and stem from different causes: they are due to phonological, morphosyntactic and processing deficits in children with SLI and to failure to access the correct form in early second language learners. This view is consistent with the Missing Surface Inflection Hypothesis (Lardière 1998; Prévost and White 2000; Haznedar 2001), which claims that EL2 children's difficulties in the production of functional morphemes is not due to a deficiency at the level of syntactic competence but rather to a problem arising at the interface between syntax and morphology. EL2 children, in fact, do not display problems in detecting errors in grammaticality judgement task, demonstrating that the underlying representations are fully specified (Ionin and Wexler 2002; Prévost and White 2000). Difficulties arise instead in production, when children have to access the morpheme that expresses a certain bundle of grammatical features (Lardiere 2009): as proposed by Guasti (in press) it seems that EL2 children sometimes fail to rapidly access their morphological system and to utter the relevant morpheme, with the consequence that inflectional features fail to be morphological expressed.

Conclusions

The main goal of our research was to examine EL2 children's performance with two clinical markers for SLI, namely the production of direct object clitics and the repetition of nonwords, in order to find similarities and differences between the two populations.

The results we obtained suggest that EL2 children have a linguistic profile which is qualitatively and quantitatively different from that typically show by children with SLI, both

in clitic production and in nonword repetition. Specifically, even though the production of clitic pronouns is problematic for EL2 as it is for children with SLI, the typologies of errors committed by the two groups are different: omission versus substitution or irrelevant production, depending on the level of general linguistic competence.

With regard to the second clinical marker for SLI we examined, the repetition of nonwords, we found that EL2 children do not exhibit difficulties and that their performance is similar to that shown by unimpaired monolingual children.

These results open up a promising line of investigation for future research, as we expect that an EL2/bilingual preschool child who actually suffers from SLI will present a remarkably low performance in nonword repetition tasks and a high omission rate in clitic production tasks. It would then be very interesting to test this prediction, comparing the performance of monolingual unimpaired children, monolingual SLI children, EL2/bilingual unimpaired children and EL2/bilingual children who have already received a diagnosis of SLI on both clinical markers. A limit of our study is that all EL2 subjects tested have a clitic system in their L1; future research could test clitic production in EL2 children whose L1 does not have clitics, in order to analyse if the typology of the first language influences their ability to produce clitics.

References

Arosio, F., Branchini, C., Barbieri, L. & Guasti, M.T. (2014). Failure to produce direct object clitic pronouns as a clinical marker of SLI in school-aged Italian speaking children. *Clinical Linguistics and Phonetics,* 9, 639-663.

Bedore, L., & Peña, E. (2008). Assessment of bilingual children for identification of language impairment: Current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism,* 11, 1–29.

Belletti, A. & Hamann, C. (2004). On the L2/bilingual acquisition of French by two young children with different source languages. In P. Prévost & D. J. Paradis (Ed.), *The acquisition of French in different contexts: Focus on functional categories* (pp. 147-174). John Benjamins.

Bialystok, E., Luk, G., Peets, K. F. & Yang S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, 13, 525-531.

Bishop, D. V. M., North, T., & Donlan, C. (1996). Nonword repetition as a phenotypic marker for inherited language impairment: Evidence from a twin study. *Journal of Child Psychology and Psychiatry*, 37, 391-342.

Bortolini, U., Caselli, M.C., Deevy, P. & Leonard, L.B, (2002). Specific language impairment in Italian: first steps in the search of a clinical marker. *International Journal of Language and Communication Disorders*, 37, 77–93.

Bortolini, U., Arfé, B., Caselli, M. C., Degasperi, L., Deevy, P. & Leonard, L.B. (2006). Clinical marker for specific language impairment in Italian: the contribution of clitics and non-word repetition. *International Journal of Language and Communication Disorders*, 41, 6, 695-712.

Caprin, C., & Guasti, M.T. (2009). The acquisition of morphosyntax in Italian: A cross-sectional study. *Applied psycholinguistics*, 30, 1, 23-52.

Casalini, C., Brizzolara, D., Chilosi, A., Cipriani, P., Marcolini, S.,Pecini, C., Roncoli, S. & Burani,C., (2007). Non-Word Repetition in Children with Specific Language Impairment: A Deficit in Phonological Working Memory or in Long-Term Verbal Knowledge? *Cortex*, 43, 769- 776.

Cecchetto, C., Di Domenico, A., Garraffa, M. & Papagno, C., (2012). *Comprendo. Batteria per la Comprensione di frasi negli adulti*. Raffaello Cortina Editore, Milano, Italy.

Coady, J. A. & Evans, J. L. (2008). Uses and interpretation of non-word repetition tasks in children with and without language impairments (SLI). *International Journal of Language and Communication Disorders*, 43, 1-40.

Conti-Ramsden, G. (2003). Processing and linguistic markers in young children with specific language impariment (SLI). *Journal of Speech, Language and Hearing Research*, 46, 1029-1037.

Cornoldi, C., Miato, L., Molin, A. & Poli, S. (2009). *PRCR-2. Prove di Prerequisito per la Diagnosi delle Difficoltà di Lettura e Scrittura.* Giunti, Organizzazioni Speciali, Florence, Italy.

Dispaldro, M., Benelli, B., Marcolini, S. & Stella, G. (2009). Real-word repetition as a predictor of grammatical competence in Italian children with typical language development. *International Journal of Language and Communication Disorders*, 44, 941-961.

Dispaldro, M.,Leonard, L.B. & Deevy, P. (2011). Real-Word and Nonword Repetition in Italian-Speaking Children With Specific Language Impairment: A Study of Diagnostic Accuracy. *Journal of Speech, Language and Hearing Research*, 56, 323-336.

Girbau, D. & Schwartz, R. G. (2008). Phonological working memory in Spanish-English bilingual children with and without Specific Language Impairment. *Journal of Communication Disorders*, 41, 124-145.

Grüter, T. (2005). Comprehension and Production of French Object Clitics by Child Second Language Learners and Children with Specific Language Impairment. *Applied Psycholinguistics*, 26, 363-391.

Guasti, M. T. (1993/1994). Verb syntax in Italian child grammar: Finite and nonfinite verbs. *Language Acquisition*, 3, 1–40.

Guasti, M. T., Maggioni, A. & Vernice M. (2013). Disturbi specifici del linguaggio, bilinguismo e acquisizione di L2. In *Nuovi contesti d'acquisizione e insegnamento: l'italiano nelle realtà plurilingui. Atti del Convegno CIS 2012 –Bergamo.* vol. 9, Guerra Edizioni, Perugia.

Guasti, M. T. (in press). *Language acquisition. The growth of grammar (2nd edition)*. Cambridge: MIT Press.

Håkansson, G. (2001). Tense morphology and verb-second in Swedish L1 children, L2 children and children with SLI. *Bilingualism: Language and Cognition*, 4, 85-99.

Håkansson, G. & Nettelbladt, U. (1993). Developmental sequences in L1 (normal and impaired) and L2 acquisition of Swedish. *International Journal of Applied Linguistics*, 3, 131-157.

Haznedar, B. (2001). The acquisition of the IP system in child L2 acquisition. *Studies in Second Language Acquisition*, 23, 1-39.

Ionin, Tania & Kenneth Wexler (2002). Why is 'is' easier than '-s'?: acquisition of tense/agreement morphology by child second language learners of English. *Second Language Research*, 18, 95-136

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards mixed models. *Journal of Memory and Language*, 59, 434-446.

Kohnert, K., Windsor, J. & Kim, D. (2006). Do language-based processing tasks separate children with Language Impairment from typical bilinguals? *Learning Disabilities Research and Practice*, 21, 19-29.

Lardiere, D. (1998). Dissociating syntax from morphology in a divergent end-state grammar. *Second Language Research*, 14, 359-75.

Lardiere, D. (2009). Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research*, 25, 173-227.

Leonard, L. (1998). *Children with specific language impairment*. Cambridge, MA: MIT Press.

Leonard, L. B., & Dispaldro, M. (2013). The Effects of Production Demands on Grammatical Weaknesses in Specific Language Impairment: The Case of Clitic Pronouns in Italian. *Journal of Speech, Language, and Hearing Research*, 56, 1272–1286.

Leonini, C. (2006). *The acquisition of object clitics and definite articles: Evidence from Italian as L2 and L1*. Unpublished doctoral disseration, Università degli studi di Siena.

Messer, M. H., Leseman, P. P. M., Boom, J. & Mayo, A. Y. (2010). Phonotactic probability effect in nonword recall and its relationship with vocabulary in monolingual and bilingual preschoolers. *Journal of Experimental Child Psychology* 105, 306-323.

Moscati, V. & Tedeschi, R. (2009). The Delay of Italian Past Participle Agreement.' In BUCLD 33: *Proceedings of the 33rd annual Boston University Conference on Language Development. Cascadilla Press.*

Oller, D. K., Pearson B. Z. & Cobo-Lewis, A.B. (2007). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics*, 28, 191–230.

Paradis, J. & Crago, M. (2000). Tense and temporality: similarities and differences between language-impaired and second-language children. *Journal of Speech, Language and Hearing Research*, 43, 834-847.

Paradis, J., Rice, M., Crago, M., & Marquis, J. (2008). The acquisition of tense in English: Distinguishing child L2 from L1 and SLI. *Applied Psycholinguistics*, 29, 1–34.

Pizzuto, E. & Caselli, M. C. (1992). The acquisition of Italian morphology: implications for models of language e development. *Journal of Child Language*, 19, 491–557.

Prévost, P. & White, L. (2000). Missing surface inflection or impairment in second language? Evidence from Tense and Agreement. *Second Language Research*, 16, 103-133.

Raven, J., Court, J.H. & Raven, J.C. (1998). *Raven Manual, Section 1 (General overview) and Section 2 (Coloured Progressive Matrices)*. Oxford, UK: Oxford Psychologist Press.

Rice, M. L. (2004). Growth models of developmental language disorders. In M. L. Rice & S. Warren (Ed.), *Developmental language disorders: From phenotypes to aetiologies*. Mahwah, NJ: Erlbaum.

Roodenrys, S. & Stokes, J. (2001). Serial recall and nonword repetition in reading disabled children. *Reading and Writing: an interdisciplinary Journal*, 14, 379-394.

Santoro, M. (2007). Second language acquisition of Italian accusative and dative clitics. *Second Language Research*, 23, 37-50.

Schaeffer, J. (2000). *The Acquisition of Direct Object Scrambling and Clitic Placement*. Amsterdam/Philadelphia: John Benjamins.

Serratrice L, Sorace A. & Paoli, S. (2004). Crosslinguistic influence at the syntax-pragmatics interface: subjects and objects in Italian-English bilingual and monolingual acquisition. *Bilingualism Language and Cognition*, 7, 183-205.

Snowling, M. J. (1981). Phonemic deficits in developmental dyslexia. *Psychological Research* 43, 219-234.

Snyder, W., Hyams, N. & Crisma, P. (1995). Romance Auxiliary Selection with Reflexive Clitics: Evidence of early knowledge of Unaccusativity. In E. Clark (Ed.) *Proceedings of the Twenty-sixth Annual Child Language Research Forum*. Stanford, CA: CSLI.

Sorace, A., Serratrice, L. Filiaci, F. & Baldo, M. (2009). Discourse conditions on subject pronoun realization: testing the linguistic intuitions of older bilingual children. *Lingua,*119, 460-477.

Stella, G., Pizzioli, C. & Tressoldi, P. (2000). *PEABODY – Test di vocabolario recettivo – P.P.V.T. – R. Peabody Picture Vocabulary Test – Revised*. Torino, Omega Edizioni.

Tamburelli, M., Sanoudaki, E., Jones, G. & Sowinska, M. (2014). Acceleration in the bilingual acquisition of phonological structure: Evidence from Polish–English bilingual children. *Bilingualism: Language and Cognition*, available on CJO2014. (doi:10.1017/S1366728914000716).

Tedeschi, R. (2006). The acquisition of object clitics in Italian: data from an elicited production task. *Annali online di Ferrara, Lettere,* 2, 31-42.

Unsworth, S., Argyri, F., Cornips L., Hulk A., Sorace, A. & Tsimpli, I (2012). The role of age of onset and input in early child bilingualism in Greek and Dutch. *Applied Psycholinguistics,* 33, 1-41.

Vender, M. (to appear). Disentangling Dylsexia: Phonological and Processing Impairment in Developmental Dyslexia. Frankfurt: Peter Lang.

Vernice, M, Arosio, F., Branchini, C., Barbieri, L., Roncaglione, E., Carravieri, E., Van der Lely, H. K. J., & Guasti, M. T. (2013). CLAD-ITA GAPS: un test di screening delle abilità fonologiche e morfo-sintattiche in bambini con un Disturbo Specifico del Linguaggio. *Psicologia Clinica dello Sviluppo*, 2, 291-314.

Windsor, J., Kohnert, K., Lobitz, K., & Pham, G. (2010). Cross-language nonword repetition by bilingual and monolingual children. *American Journal of Speech-Language Pathology*, 19, 298-310.

Zesiger, P. Chillier-Zesiger, L., Arabatzi, M., Baranzini, L., Cronel-Ohayon, S., Franck, J., Frauenfelder, U.H., Hamann, C., & Rizzi, L. (2010). The acquisition and mastery of French pronouns in normal children. *Applied Psycholinguistics*, 31, 571-603.

Table 1. Descriptive data concerning 3 groups of EL2 children, Albanian (ALB), Arabic (ARA) and Romanian (RUME), and one group of Monolingual control children (MON) matched for chronological age. For each group, we reported means and standard deviation (SD) of the participants' age at the moment of testing (AGE), Age at First Exposure to Italian (AFE), Quantity of Exposure (QE), Traditional Length of Exposure (TLE), Cumulative Length of Exposure (CLE), raw scores obtained in the grammatical comprehension test Comprendo, in the PPVT-R, in nonword repetition and in the Raven test.

	AGE (SD)	AFE (SD)	QE (SD)	TLE (SD)	CLE (SD)	COMPRENDO (SD)	PPVT-R (SD)	RAVEN (SD)
ALB	4.98	1.22	0.66	3.83	2.10	14.55	77.13	14.90
n°40	(0.70)	(1.47)	(0.13)	(1.45)	(0.84)	(3.25)	(9.68)	(3.99)
ARA	4.78	1.39	0.55	3.46	1.59	12.78	71.19	13.70
n°37	(0.51)	(1.40)	(0.15)	(1.48)	(0.69)	(3.29)	(8.26)	(2.90)
RUME	4.77	1.58	0.58	3.38	1.89	14.76	79.45	15.74
n°38	(1.00)	(1.43)	(0.19)	(1.50)	(1.21)	(3.30)	(11.72)	(3.54)
MON	4.83	-	-	-	-	15.62	88.03	15.13
n°39	(0.52)					(2.67)	(12.33)	(2.98)

Table 2. Mean percentages of responses (and *SDs*) for each participant group in the Clitic Pruduction Task

	TARGET (SD)	INCORRECT_CL (SD)	NP (SD)	OMISSION (SD)	OTHER (SD)
ALB	0.48	0.33	0.06	0.04	0.09
	(0.50)	(0.47)	(0.24)	(0.19)	(0.28)
ARA	0.29	0.24	0.09	0.09	0.28
	(0.45)	(0.42)	(0.28)	(0.29)	(0.44)
RUME	0.51	0.29	0.06	0.08	0.06
	(0.50)	(0.45)	(0.23)	(0.26)	(0.24)
MON	0.73	0.16	0.06	0.03	0.02
	(0.44)	(0.36)	(0.22)	(0.17)	(0.13)

Table 3 Summary of the fixed effects in the mixed logit models for monolingual and EL2 children (N= 1848, children=154) for target clitics, incorrect clitics and other responses.

Predictor	**Est**	**St. Err**	**Df**	**T**	**p**
Target (loglikehood = 1088)					
(Intercept)	0.23	0.10	157	2.1	< 0.01
MON vs ALB	-0.21	0.05	149	-3.96	<0.001
MON vs ARA	-0.35	0.05	149	-6.1	<0.001
MON vs RUME	-0.03	0.006	149	5.1	<0.001
ARA vs ALB	0.19	0.05	150	3.26	< 0.01
ARA vs RUME	0.22	0.06	150	3.78	< 0.01
Comprendo	0.03	0.006	149	5.10	<0.001
Note. Random effects for subjects and items had SD of 0.21 and 0.09, respectively.					
Incorrect clitic (loglikehood= -969)					
(Intercept)	0.16	0.04	49	3.7	< 0.01
MON vs ALB	0.17	0.04	150	3.6	< 0.001
MON vs RUME	0.13	0.04	150	2.7	< 0.01
ARA vs ALB	0.09	0.04	150	2.0	< 0.05
Note. Random effects for subjects and items had SD of 0.17 and 0.09, respectively					
Other (loglikehood= -126)					
(Intercept)	0.49	0.13	150	3.6	< 0.01
MON vs ARA	0.17	0.04	148	3.53	< 0.01
ARA vs ALB	-0.18	0.04	150	-4.4	< 0.001
ARA vs RUME	-0.21	0.04	150	-4.9	< 0.001
Comprendo	-0.01	0.004	148	-2.88	< 0.01
PPVT-R	-0.002	0.001	148	-1.99	< 0.05
Note. Random effects for subjects and items had SD of 0.16 and 0.04, respectively					

Table 4 Summary of the fixed effects in the mixed logit models for 115 EL2 children (N= 1380) for target and incorrect clitics.

Predictor	**Est**	**St. Err**	**Df**	**t**	**p**
Target (loglikehood = -821)					
(Intercept)	2.399e-01	7.110e-02	8.871e+01	3.3	< 0.01
Group= ARA	-1.329e-01	5.724e-02	1.110e+02	-2.3	0.02
CLE	4.785e-03	9.775e-04	1.110e+02	4.896	<0.01
Note. Random effects for subjects and items had SD of 0.21 and 0.011, respectively.					
Incorrect clitic (loglikehood= -969)					
(Intercept)	2.022e-01	5.517e-02	1.176e+02	3.665	< 0.01
Group= ARA	1.613e-01	4.923e-02	1.110e+02	3.276	< 0.01
CLE	0.13	8.407e-04	1.110e+02	-2.678	< 0.001
Note. Random effects for subjects and items had SD of 0.19 and 0.04, respectively					

Table 5. Raw numbers of errors committed by each groups according to their type.

	Gender	**Number**	**Gender and Number**	**Case**
MON	35	13	18	4
ALB	86	31	19	17
ARA	39	12	23	20
RUME	61	33	23	3
Total	221	89	83	44

Table 6 Summary of the fixed effects in the mixed logit models for 115 EL2 children (N= 436) for gender and number error.

Predictor	Est	St. Err	Wald Z	p
Gender Error (loglikehood = -575.2)				
(Intercept)	-3.47	0.52	-6.62	< 0.001
MON vs ALB	1.22	0.29	4.15	<0.01
MON vs ARA	0.18	0.31	0.59	NS
MON vs RUME	0.79	0.30	2.61	<.01
Note. Random effects for subjects and items had SD of 0.80 and 1.55, respectively.				
Number Error (loglikehood= -278.2)				
(Intercept)	-5.73	0.87	-6.56	< 0.001
MON vs ALB	1.09	0.48	2.27	< 0.05
MON vs ARA	-0.05	0.54	-0.10	NS
MON vs RUME	1.29	0.48	2.67	<.01
Note. Random effects for subjects and items had SD of 1.17 and 2.15, respectively				

Table 7 The frequency of clitic substitution errors committed by each participant group, indicating how many times a given clitic was wrongly chosen..

	lo	***la***	***le***	***li***	**Tot**
MONO	31	23	1	10	65
ALB	46	26	3	28	103
ARA	14	18	2	19	53
RUM	56	12	0	6	74

Table 8. Number of Target RE clitics, Incorrect Clitics, Omissions and Other irrelevant responses produced by the four groups of subjects

	Target	**Incorrect clitic**	**Omissions**	**Other**
MONO	157	1	2	0
ALB	114	3	13	18
ARA	68	9	44	42
RUME	138	1	14	7

Table 9 Summary of the fixed effects in the mixed logit models for children monolingual and L2 children (N= 1848, children=154) (loglikehood=-99.2) for target RE clitics, omissions, other clitics and other responses.

Predictor	Est	St. Err	Df	t	p
Target (loglikehood = -99.2)					
(Intercept)	0.23	0.22	148	1.025	NS
MON vs ARA	-0.45	0.07	148	-5.6	<0.001
Comprendo	0.017	0.008	148	2.1	<0.05
PPVT-R	0.005	0.002	148	2.1	<0.05
Note. Random effects for subjects and items had SD of 0.29 and 0.02, respectively.					
Omission (loglikehood= -110.59)					
(Intercept)	0.012	0.03	114	0.3	NS
MON vs ARA	0.26	0.05	150	5.2	< 0.001
Note. Random effects for subjects and items had SD of 0.18 and 0.005, respectively					
Other clitics (loglikehood= 388)					
(Intercept)	6.4e-03	1.3e-02	4.3e+01	0.4	NS
MON vs ARA	4.9e-02	1.8e-02	1.5e+02	2.7	< 0.01
Note. Random effects for subjects and items had SD of 0.3 and 0.004, respectively					
Other responses (loglikehood= 5.38)					
(Intercept)	0.27	0.1	149	2.6	<.01
MON vs ARA	0.21	0.05	149	3.8	< 0.01
Comprendo	0.01	0.006	149	-2.9	<0.05
Note. Random effects for subjects and items had SD of 0.21 and 0.03, respectively					

Table 10 Z-scores of the four groups of participants in the Nonword Repetition Task

	NONWORDS (SD)
ALB	0.04 (1.10)
ARA	-0.20 (1.20)
RUME	-0.05 (1.11)
MONO	0.11 (0.69)